

Squared Nearest-Neighbor Distances in the Unit Square

Let $p_1, \dots, p_n$ be labelled points in

$$U = [0, 1]^2;$$

coincidences are allowed. Define

$$r_i = \min_{j \neq i} |p_i - p_j|, \quad S = \sum_{i=1}^n r_i^2, \quad R = \max_i r_i.$$

We use Pinelis's rectangle theorem: if a finite set of points lies in an $a \times b$ rectangle, then the sum of its squared nearest-neighbor distances is at most

$$2a^2 + 2b^2.$$

For labelled points with coincidences the same upper bound remains valid.

1 A very simple bound below 4

Theorem 1. For every $n \geq 5$,

$$S \leq \frac{15}{4} = 3.75 < 4.$$

1.1 Crude small- R disk estimate

For each i draw the closed disk D_i centered at p_i with radius $r_i/2$. The interiors of these disks are pairwise disjoint.

Every D_i lies in the $R/2$ -neighborhood of U , whose area is

$$1 + 2R + \frac{\pi}{4}R^2.$$

Hence

$$\frac{\pi}{4}S \leq 1 + 2R + \frac{\pi}{4}R^2,$$

so

$$S \leq \frac{4}{\pi} + \frac{8R}{\pi} + R^2.$$

The right side is increasing. Therefore, if

$$R \leq \frac{1}{\sqrt{2}},$$

then

$$S \leq \frac{1}{2} + \frac{4(1 + \sqrt{2})}{\pi} = 3.573872177 \dots < \frac{15}{4}.$$

1.2 An easy adjacent-quarter lemma

Lemma 1. Let A be a nonempty collection of labelled points in one quarter-square, and let q lie in an adjacent quarter-square. If $\Sigma(A)$ denotes the sum of squared nearest-neighbor distances computed inside A , then

$$\Sigma(A) + \text{dist}(q, A)^2 \leq \frac{7}{4}.$$

Proof. Scale lengths by 2. It is enough to prove the bound 7 in the 2×1 picture.

Move q horizontally to the right edge, so $q = (2, v)$. Translate A horizontally left until the left edge of its minimal bounding rectangle is $x = 0$. Write this rectangle as

$$[0, x] \times [\gamma, \delta],$$

with height $h \leq 1$.

Pinelis gives

$$\Sigma(A) \leq 2x^2 + 2h^2.$$

There is a point of A on the right side of the bounding rectangle. Its horizontal distance from q is $2 - x$, and its vertical distance is at most 1. Hence

$$\text{dist}(q, A)^2 \leq (2 - x)^2 + 1.$$

Thus

$$\Sigma(A) + \text{dist}(q, A)^2 \leq 2x^2 + 2 + (2 - x)^2 + 1 = 3x^2 - 4x + 7.$$

This quadratic is convex on $0 \leq x \leq 1$, so its maximum is at an endpoint. The endpoint values are 7 and 6. Therefore the scaled bound is 7, and scaling back gives $7/4$. $\square$

1.3 Large- R estimate for Theorem 1

Assume

$$R > \frac{1}{\sqrt{2}},$$

choose q with $r_q = R$, and divide U into four quarter-squares. The quarter containing q is a singleton because a quarter-square has diameter $1/\sqrt{2}$. Any quarter containing at least two points contributes at most 1 internally by Pinelis.

If all three other quarters are nonsingletons, pair q with either adjacent quarter and use Lemma 1; the other two quarters contribute at most 1 each. Thus

$$S \leq \frac{7}{4} + 2 = \frac{15}{4}.$$

If exactly two other quarters are nonsingletons and the third is a singleton, pair the two singletons with the two nonsingleton quarters, giving

$$S \leq 2 \cdot \frac{7}{4} = \frac{7}{2}.$$

If the third is empty, then

$$S \leq \frac{7}{4} + 1 = \frac{11}{4}.$$

If exactly one other quarter is nonsingleton, Pinelis's Case 4/5 estimate gives

$$S \leq 3.$$

Combining these bounds with Section 1.1 proves Theorem 1.

2 A sharper disk-covering estimate

We now improve the small- R part by estimating more carefully how much the pairwise disjoint disks can cover outside U . Put

$$a = \frac{R}{2}.$$

2.1 A cap below one side

Lemma 2. Let D be a disk of radius ρ centered at (x_0, y_0) , with $y_0 \geq 0$, and suppose D extends below $y = 0$. Put

$$h = \sqrt{\rho^2 - y_0^2}.$$

Then the portion of D below $y = 0$ has area at most

$$\min \left\{ \frac{\pi}{2} h^2, 2h(\rho - y_0) \right\}.$$

The cap is pointwise dominated by the lower semicircle of radius h centered at $(x_0, 0)$, and it also lies in a rectangle of width $2h$ and height $\rho - y_0$.

The diagram shows a circle C centered at $C = (x_0, y_0)$ with radius ρ . A horizontal interval $I = [x_0 - h, x_0 + h]$ is marked on the x-axis. A point Q is located on the boundary of the circle. A green line segment connects the center C to the point Q . A dashed red rectangle is drawn below the x-axis, and a dashed black arc is shown below the rectangle.

dashed: lower semicircle of radius h

dash-dot: rectangle of height $\rho - y_0$

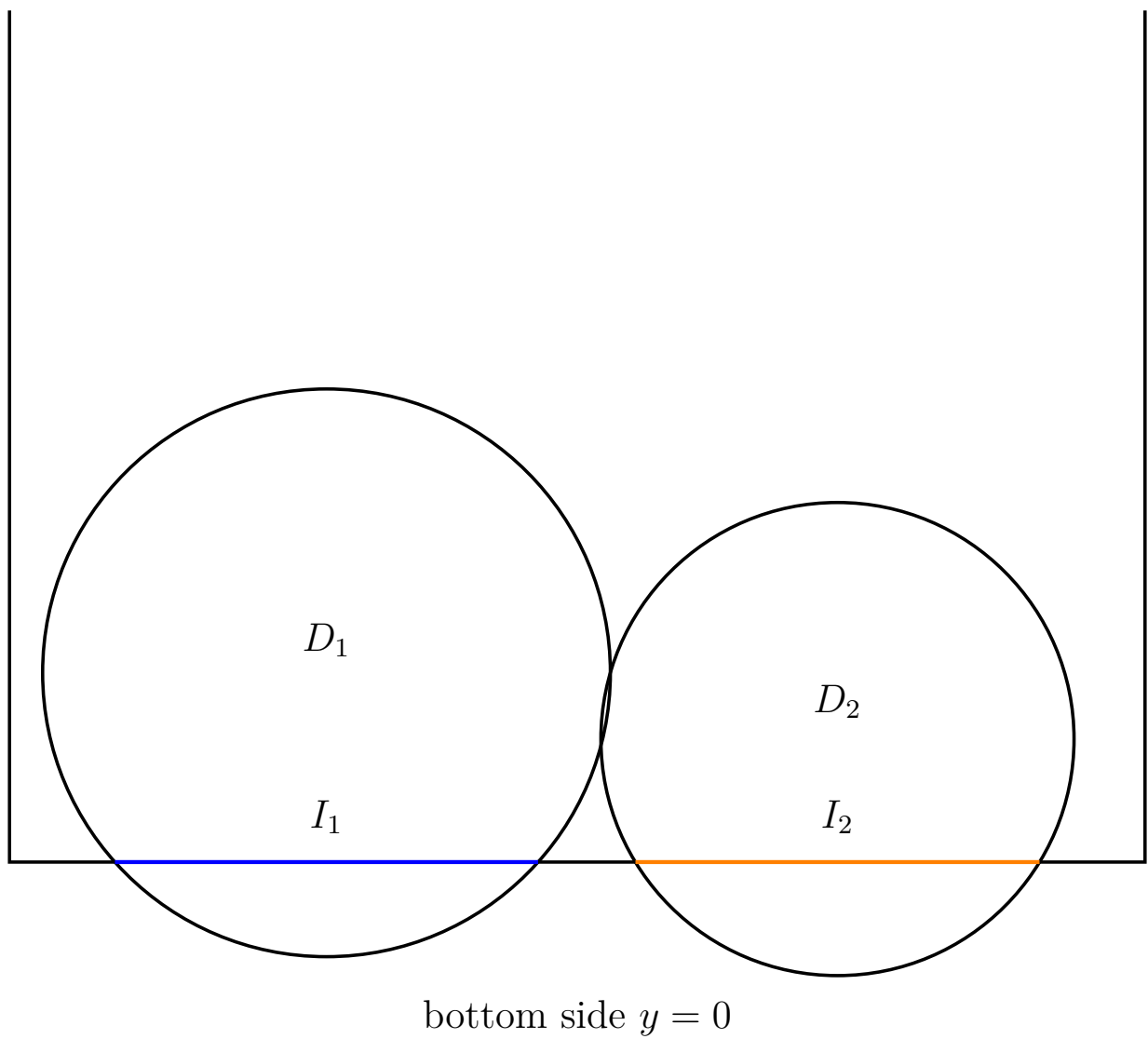
Definition. The chord interval is

$$I = [x_0 - h, x_0 + h].$$

Lemma 3. If two original disks are disjoint, then the interiors of their chord intervals are disjoint.

Indeed, an interior point common to both intervals would give a point $(x, 0)$ lying strictly inside both disks.

Figure 2: disjoint disks give disjoint chord intervals



2.2 An endpoint estimate

Define

$$t_0 = \sqrt{1 - \frac{\pi^2}{16}}, \quad e_0 = \frac{1}{2} \arcsin(t_0) - \frac{\pi}{8} t_0.$$

Lemma 4. Let $0 < h \leq a$. Consider a lower semicircle of radius h centered at $(c, 0)$, with $0 \leq c \leq h$. Let A be the area of the part with $x \geq 0$ and let

$$L = c + h.$$

Then

$$A \leq \frac{\pi a}{4} L + a^2 e_0.$$

Put $t = c/h$ and $s = h/a$. Direct integration gives

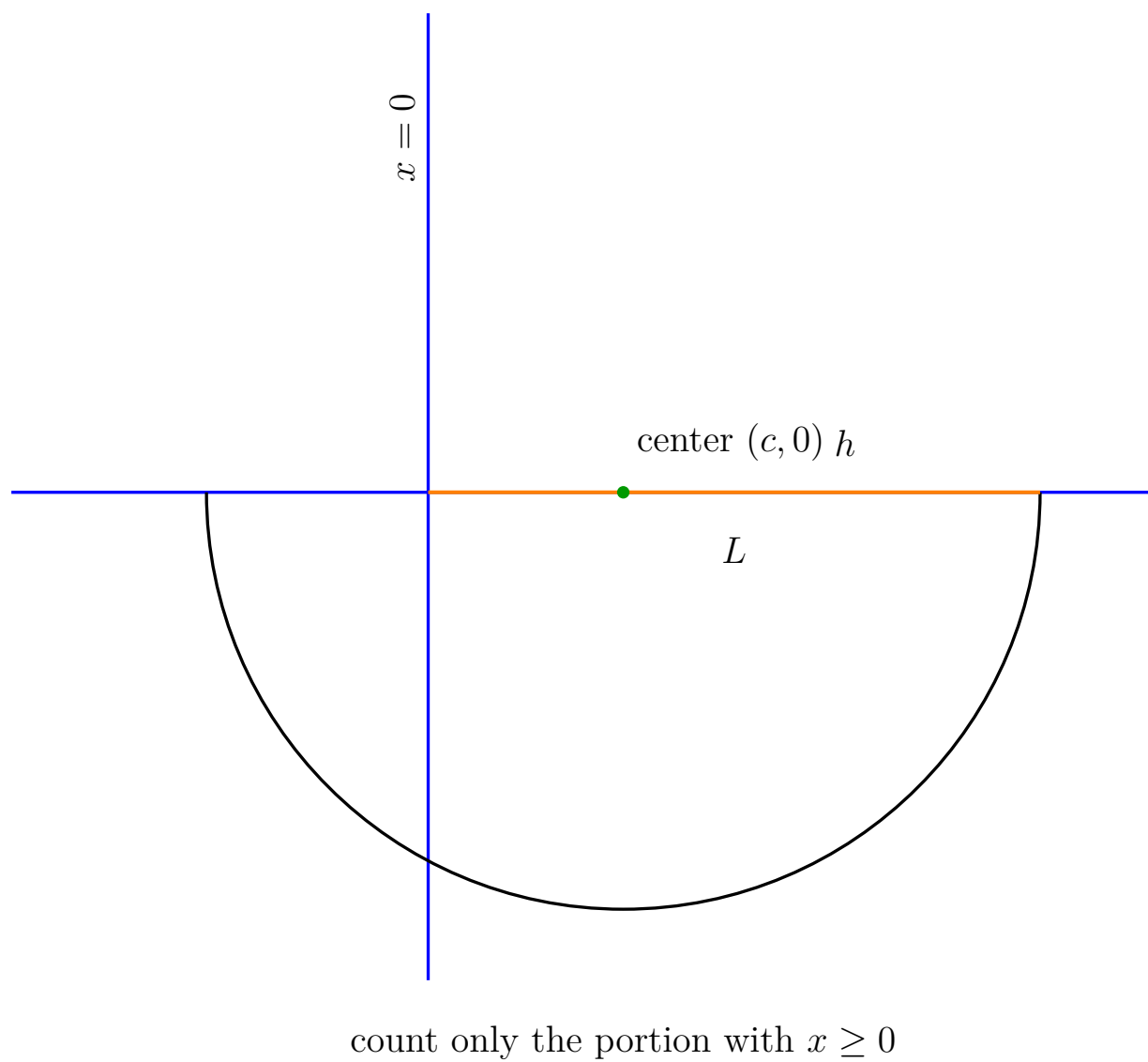
$$A = h^2 \left[\frac{\pi}{4} + \frac{1}{2} \left(t \sqrt{1 - t^2} + \arcsin t \right) \right].$$

For fixed t , the normalized excess over $(\pi a/4)L$ is a convex quadratic in s , so its maximum occurs at $s = 0$ or $s = 1$. At $s = 1$ its derivative is

$$\sqrt{1 - t^2} - \frac{\pi}{4},$$

which is maximized at $t = t_0$.

Figure 3: the endpoint case in Lemma 4



2.3 One side strip

Lemma 5. In the bottom strip

$$B = \{(x, y) : 0 \leq x \leq 1, y < 0\},$$

the total area covered by the original disks is at most

$$M(a) = \max \left\{ \frac{\pi a}{4} + 2a^2 e_0, \frac{\pi a^2}{2} \right\}.$$

Apply Lemma 2 to the protruding caps. The chord intervals have disjoint interiors. For intervals inside $[0, 1]$, bound the comparison semicircle by $(\pi a/4)$ times the chord length. For at most two endpoint-crossing intervals, use Lemma 4. If one interval contains all of $[0, 1]$, the total is at most one full semicircle of radius a .

Figure 4A: original disjoint configuration

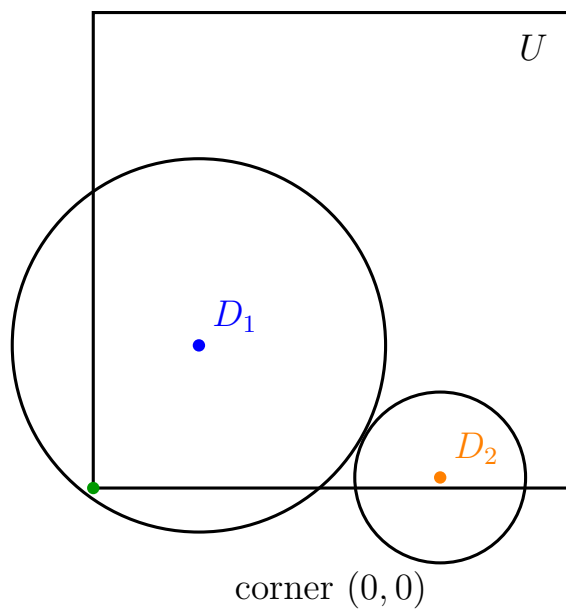
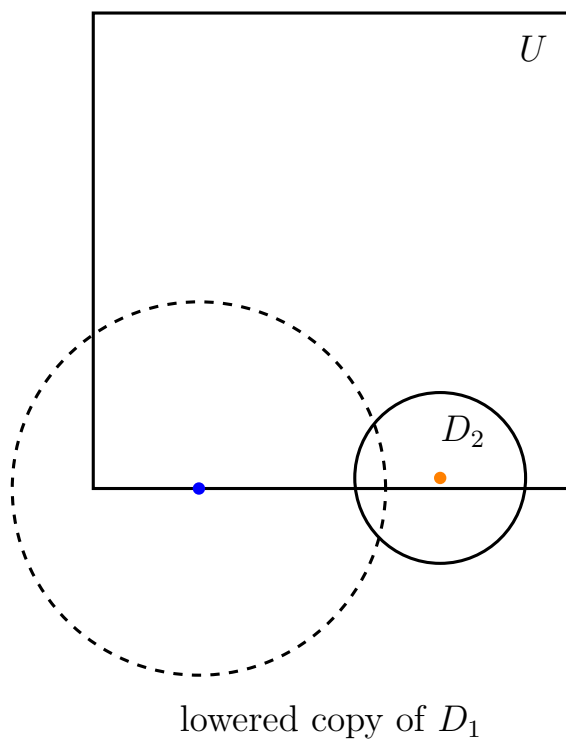


Figure 4B: lower only D_1 until its center lies on $y = 0$



2.4 The improved covering theorem

Put

$$C = \pi + 8e_0 = 3.86680427735\dots$$

Theorem 2. Let $D_1, \dots, D_n$ be pairwise disjoint disks centered in U and of radii at most $R/2$. Then

$$\sum_i \text{area}(D_i) \leq 1 + \max \left\{ \frac{\pi}{2}R + \frac{C}{4}R^2, \frac{3\pi}{4}R^2 \right\}.$$

The part inside U has area at most 1. Each of the four corner quadrants contributes at most a quarter-disk of radius a , for a total of πa^2 . Each side strip contributes at most $M(a)$. Thus the total is at most

$$1 + \pi a^2 + 4M(a);$$

substituting $a = R/2$ gives the theorem.

For the nearest-neighbor disks,

$$\sum_i \text{area}(D_i) = \frac{\pi}{4}S,$$

hence

$$S \leq \frac{4}{\pi} + \max \left\{ 2R + \frac{C}{\pi}R^2, 3R^2 \right\}.$$

At

$$R = \frac{1}{\sqrt{2}},$$

this gives

$$S \leq 3.3028741218\dots$$

3 The sharp adjacent-quarter lemma

Lemma 6. Let A be a finite nonempty collection of labelled points in one quarter-square, and let q lie in an adjacent quarter-square. Then

$$\Sigma(A) + \text{dist}(q, A)^2 \leq \frac{25}{16}.$$

The constant $25/16$ is sharp.

Proof. Scale lengths by 2. Move q horizontally to the right edge, so

$$q = (2, v).$$

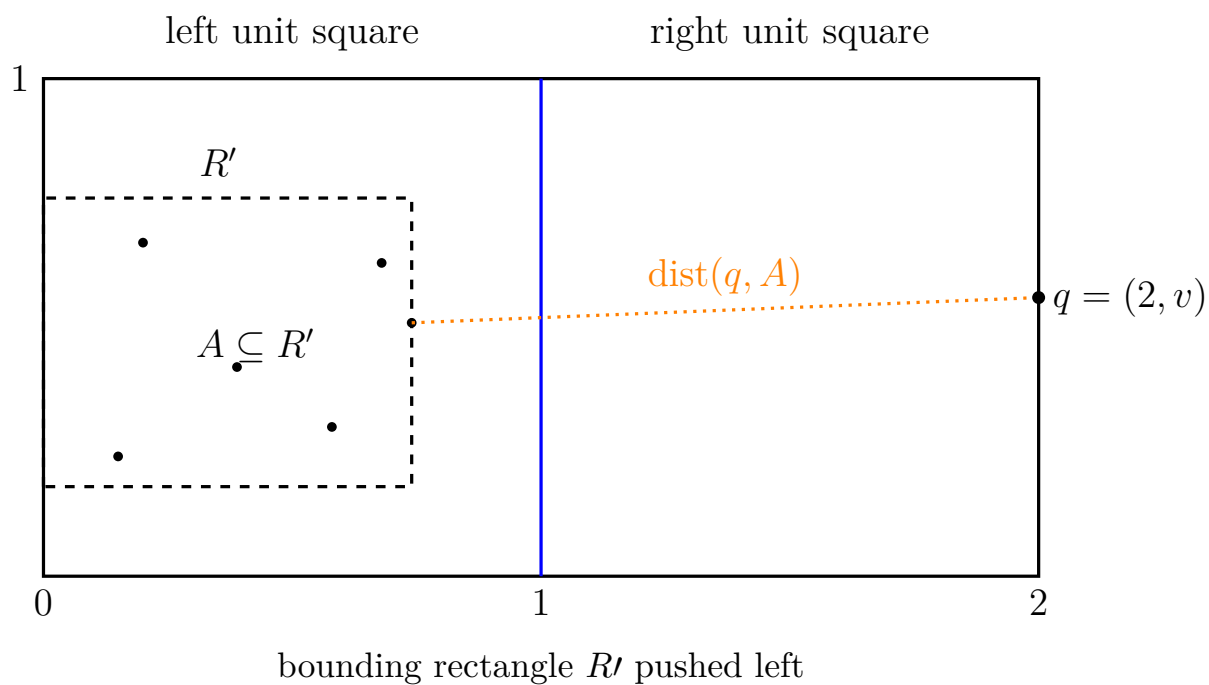
Translate A horizontally left until the left edge of its minimal bounding rectangle is $x = 0$. Write that rectangle as

$$[0, x] \times [\gamma, \delta],$$

with height

$$h = \delta - \gamma \leq 1.$$

Figure 5: the scaled adjacent-quarter geometry



By Pinelis,

$$\Sigma(A) \leq 2x^2 + 2h^2.$$

Put

$$a = 2 - x, \quad \Delta = 4 - a^2 = 4x - x^2.$$

Let

$$s = \min\{(v - \gamma)^2, (v - \delta)^2\}, \quad t = \max\{(v - \gamma)^2, (v - \delta)^2\}.$$

A right-side point gives

$$\text{dist}(q, A)^2 \leq a^2 + t,$$

while the nearer top/bottom-side point gives

$$\text{dist}(q, A)^2 \leq a^2 + \Delta + s.$$

Hence

$$\text{dist}(q, A)^2 \leq a^2 + \min\{t, s + \Delta\}.$$

The needed vertical estimate is

$$2h^2 + \min\{t, s + \Delta\} \leq \frac{9}{4} + \frac{3}{4}\Delta \quad (0 \leq \Delta \leq 1),$$

and

$$2h^2 + \min\{t, s + \Delta\} \leq 3 \quad (\Delta \geq 1).$$

If v lies between γ and δ , let $p \leq q_0$ be the two vertical distances. Then

$$p + q_0 = h \leq 1$$

and

$$\min\{q_0^2, p^2 + \Delta\} \leq \frac{(1 + \Delta)^2}{4}.$$

If v lies outside the interval, let p be the nearer vertical distance. Then

$$p + h \leq 1$$

and

$$2h^2 + p^2 + \Delta \leq 2 + \Delta.$$

These give the displayed estimate.

When $0 \leq \Delta \leq 1$, equivalently

$$0 \leq x \leq 2 - \sqrt{3},$$

the bound becomes

$$\frac{25}{4} - x + \frac{9}{4}x^2 \leq \frac{25}{4}.$$

When $\Delta \geq 1$, the bound becomes

$$7 - 4x + 3x^2.$$

By convexity its maximum on $[2 - \sqrt{3}, 1]$ is at an endpoint, with values 6 and

$$20 - 8\sqrt{3} < \frac{25}{4}.$$

Scaling back proves the lemma.

Sharpness comes from

$$A = \{(0, 0), (0, 1)\}, \quad q = (2, 1/2)$$

in the scaled picture. □

4 The Global Theorem

Theorem 3. Let $p_1, \dots, p_n$ be labelled points in the unit square, with $n \geq 5$. Let r_i be the distance from p_i to its nearest other point, and put

$$S = \sum_i r_i^2.$$

Then

$$S \leq \frac{57}{16} = 3.5625.$$

Proof. Put

$$R = \max_i r_i.$$

We divide the proof into two cases.

Case 1: $R \leq 1/\sqrt{2}$. For each i draw the disk D_i centered at p_i with radius $r_i/2$. These disks have pairwise disjoint interiors, and

$$\sum_i \text{area}(D_i) = \frac{\pi}{4} S.$$

Theorem 2 gives, for every $R \geq 0$,

$$\sum_i \text{area}(D_i) \leq 1 + \max \left\{ \frac{\pi}{2} R + \frac{C}{4} R^2, \frac{3\pi}{4} R^2 \right\}.$$

The resulting bound for S is increasing on $0 \leq R \leq 1/\sqrt{2}$, and at $R = 1/\sqrt{2}$ it is

$$S \leq 3.3028741218 \dots < \frac{57}{16}.$$

Therefore the theorem holds whenever $R \leq 1/\sqrt{2}$.

Case 2: $R > 1/\sqrt{2}$. Choose q with $r_q = R$, and divide the unit square into four quarter-squares of side $1/2$. A quarter-square has diameter $1/\sqrt{2}$, so no other labelled point can lie in the same quarter-square as q . Thus the quarter containing q is a singleton.

For any other quarter-square Q containing at least two points, let $\Sigma(Q)$ be the sum of the squared nearest-neighbor distances computed using only the points of Q . The nearest-neighbor distance of a point in the full configuration is no larger than its nearest-neighbor distance inside Q , so the contribution of Q to S is at most $\Sigma(Q)$. By Pinelis's rectangle theorem,

$$\Sigma(Q) \leq 2 \left(\frac{1}{2}\right)^2 + 2 \left(\frac{1}{2}\right)^2 = 1.$$

We also use Lemma 6. If A lies in a quarter-square adjacent to q 's quarter, then

$$\Sigma(A) + \text{dist}(q, A)^2 \leq \frac{25}{16}.$$

Since

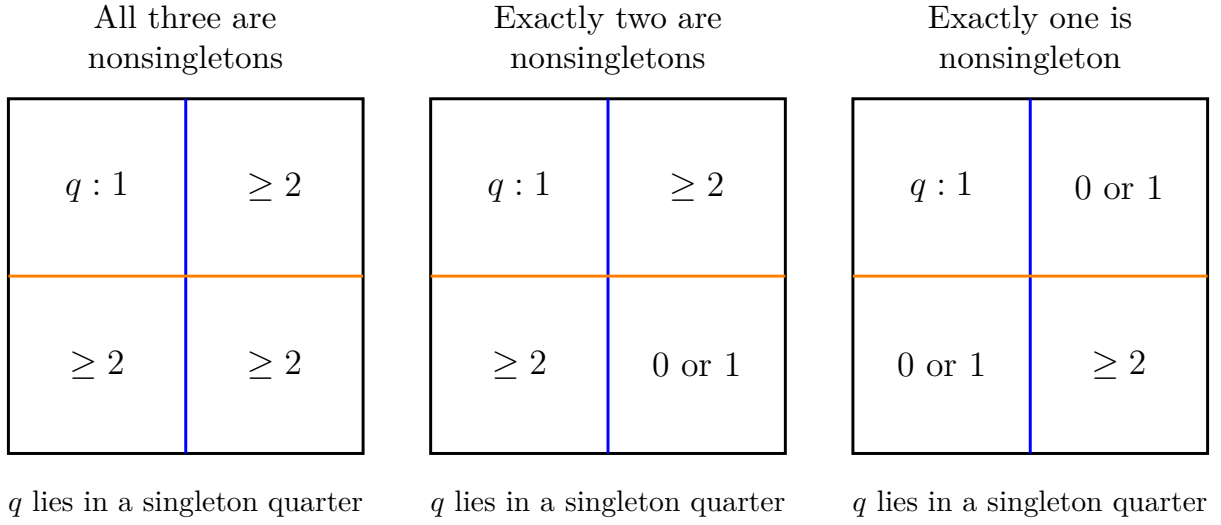
$$r_q \leq \text{dist}(q, A),$$

we may use the weaker but more convenient form

$$\Sigma(A) + r_q^2 \leq \frac{25}{16}.$$

There are three quarter-squares other than q 's singleton quarter. We classify the configuration by how many of these three quarters are nonsingletons.

Figure 6: the large- R occupancy cases



All three are nonsingletons. Choose either nonsingleton quarter adjacent to q 's quarter and call its point set A . Lemma 6 gives

$$\Sigma(A) + r_q^2 \leq \frac{25}{16}.$$

Each of the remaining two nonsingleton quarters contributes at most 1. Hence

$$S \leq \frac{25}{16} + 1 + 1 = \frac{57}{16}.$$

Exactly two are nonsingletons. If the third quarter is empty, at least one nonsingleton quarter is adjacent to q 's quarter. Pair it with q using Lemma 6; the other nonsingleton quarter contributes at most 1. Thus

$$S \leq \frac{25}{16} + 1 = \frac{41}{16}.$$

If the third quarter is a singleton, there are two singleton quarters and two nonsingleton quarters. They can be paired so that each singleton is adjacent to one nonsingleton quarter. Applying Lemma 6 twice gives

$$S \leq 2 \left(\frac{25}{16} \right) = \frac{25}{8}.$$

Exactly one is nonsingleton. The other two quarters are singleton or empty. This is Pinelis's Case 4, with Case 5 allowing empty quarters. His stronger intermediate estimate gives

$$S \leq \frac{3}{2}(a^2 + b^2)$$

for an $a \times b$ rectangle. For the unit square,

$$S \leq 3.$$

Thus in every large- R configuration,

$$S \leq \frac{57}{16}.$$

Combining this with Case 1 proves Theorem 3. □

5 Conditional improvement from a 21/16 lemma

Suppose the following strengthening is true: whenever A contains at least three labelled points in a quarter-square and q lies in an adjacent quarter-square,

$$\Sigma(A) + \text{dist}(q, A)^2 \leq \frac{21}{16}.$$

Then for $n \geq 10$ one obtains a better bound. If an adjacent quarter contains at least three points, the conditional lemma together with the two remaining quarter bounds gives at most

$$\frac{21}{16} + 1 + 1 = \frac{53}{16} = 3.3125.$$

Otherwise the potentially worst occupancy is

$$1 + 2 + 2 + m.$$

For $n \geq 10$ we have $m \geq 5$. Pair q with one 2-point adjacent quarter using Lemma 6, bound the other 2-point quarter by 1, and apply the proved 57/16 theorem to the diagonal m -point quarter after scaling that quarter to a unit square. Its contribution is at most

$$\frac{1}{4} \cdot \frac{57}{16} = \frac{57}{64}.$$

Therefore, conditionally,

$$S \leq \frac{25}{16} + 1 + \frac{57}{64} = \frac{221}{64} = 3.453125 \quad (n \geq 10).$$

For $n = 8$, the remaining pattern $1 + 2 + 2 + 3$ gives

$$S \leq \frac{25}{16} + 1 + \frac{13}{16} = \frac{27}{8} = 3.375.$$

For $n = 7$ and $n = 9$, the patterns $1 + 2 + 2 + 2$ and $1 + 2 + 2 + 4$ are not eliminated merely by the conditional lemma.

Reference

I. Pinelis, “Exact upper bound on the sum of squared nearest-neighbor distances between points in a rectangle,” *Mathematical Inequalities & Applications* **23** (2020), 663–670. DOI: 10.7153/mia-2020-23-54; arXiv:1904.11427.